

UWU/CST/21/050 - W.D.V. Asini Weerasinghe

## Assignment on Memory Management

CST 262-2

- 01) Compare paging with Segmentation regarding the amount of memory required by the address translation structures to convert virtual address to physical addresses.

### Paging

- \* page Table: maps virtual pages to physical frames.
- \* Memory Required: Depends on the number of entries, determined by the virtual address space and page size

Ex:- for a 32 bit address space with 4kB pages

- Number of entries:  $2^{20}$
- Entry Size: 4 bytes
- Total memory: 4MB

### Segmentation

- \* Segment table: Maps segment identifiers to base addresses & limit.
- \* Memory required: Depends on the number of segments & entry size.

Ex:- For up to 16 segments.

- Number of entries: 16
- Entry Size 8 bytes
- Total memory: 128 bytes.

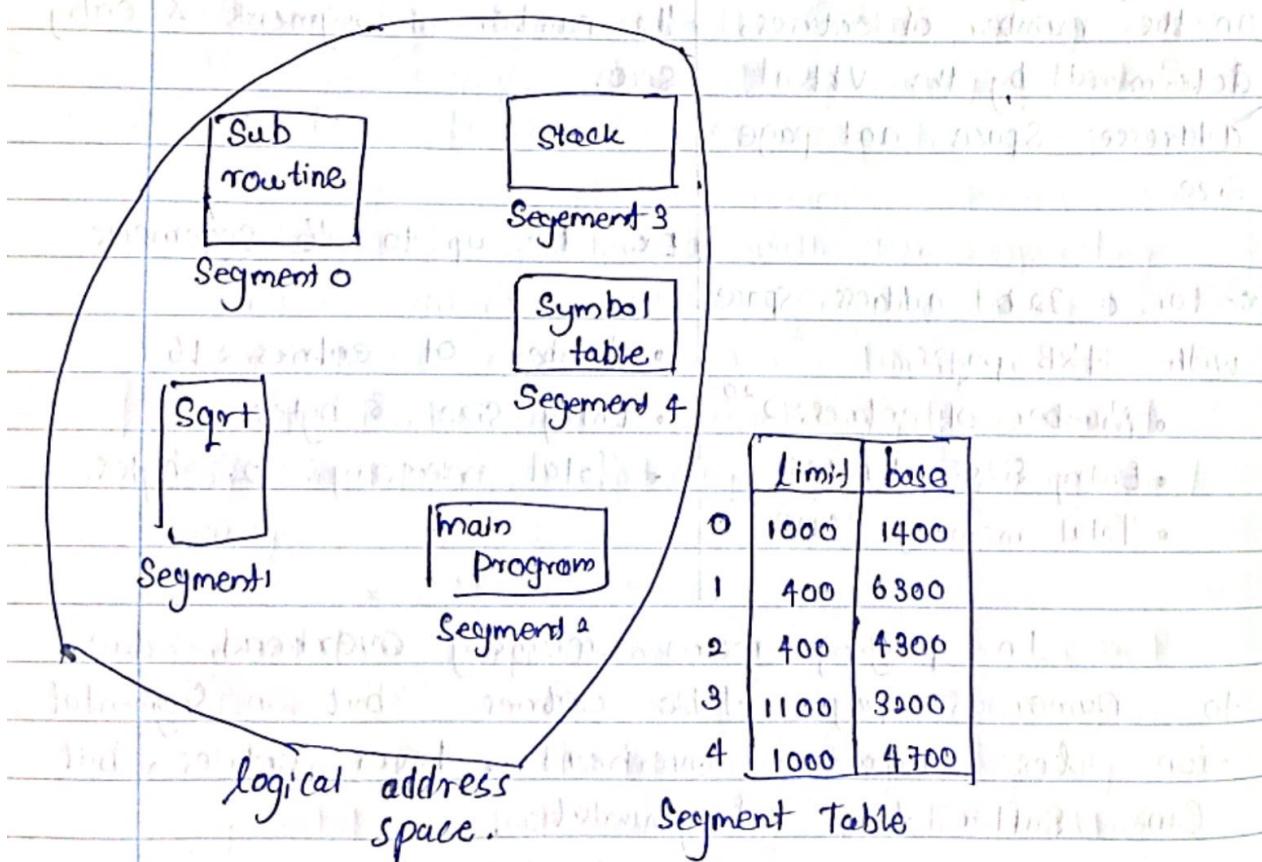
In paging more memory overhead due to numerous page table entries but in Segmentation less memory overhead, fewer entries, but can suffer from fragmentation.

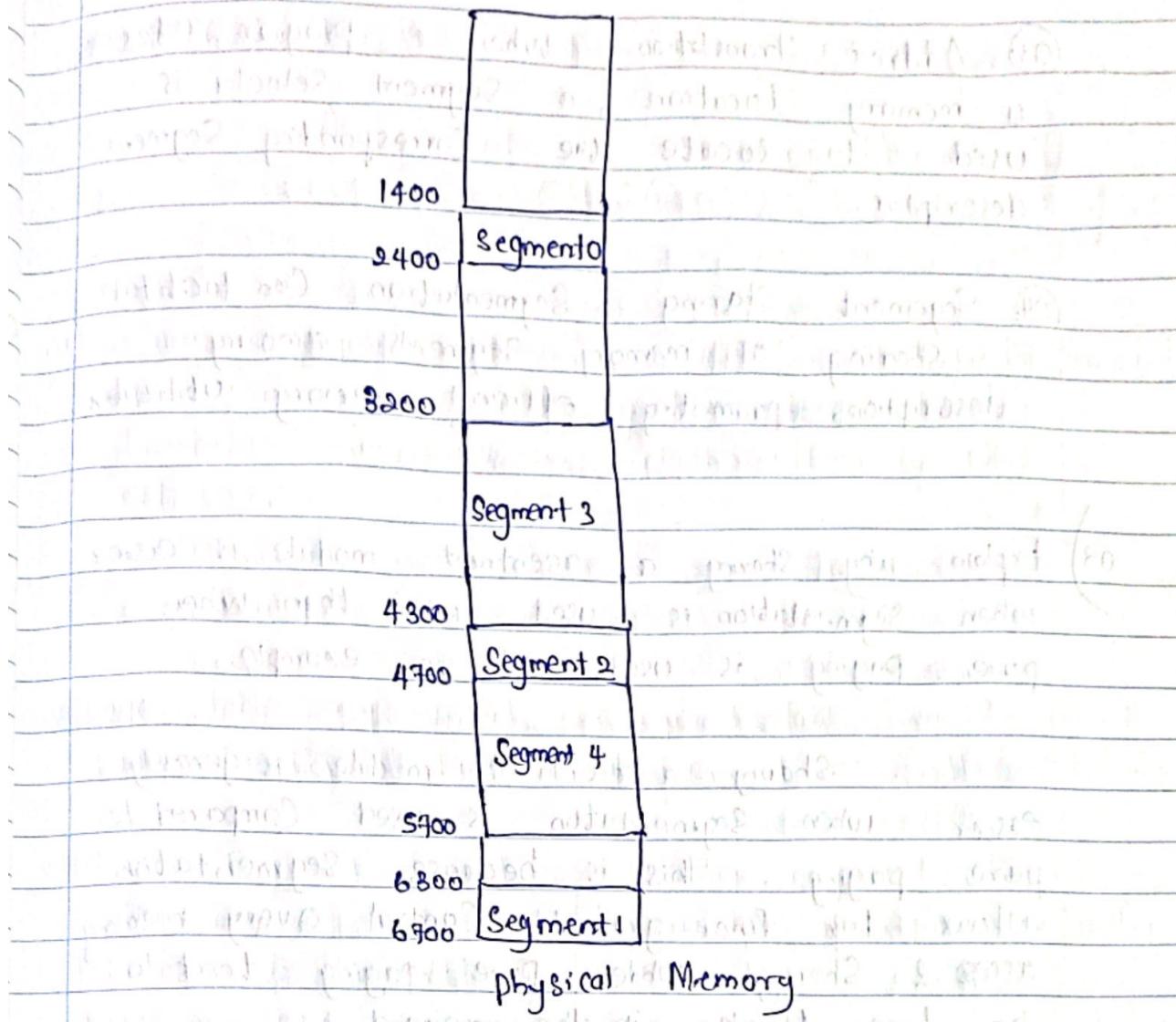
Paging, Fragmentation - paging can lead to internal fragmentation because pages are typically of a fixed size. This means that some memory within a page may remain unused, resulting in wasted space.

Segmentation, Fragmentation - Segmentation can lead to external fragmentation because segments are not of fixed size.

As segments are allocated, gaps may appear in memory that are too small to hold a new segment but too large to be used efficiently.

- Q2) Draw the diagram of the Segmentation memory management scheme & explain its principle.





### Principles of Segmentation Management.

- ① Segmentation Units - Memory is divided into segments, where each segment represents a logical unit such as code segment, data segment or stack segment.
- ② Segment Descriptors - Each segment is associated with a segment descriptor, which contains information about the segment's location, size & access rights.

(03) Address Translation - When a program refers to a memory location, a Segment Selector is used to locate the corresponding Segment descriptor.

(04) Segment Sharing - Segmentation can facilitate sharing of memory segments among processes, promoting efficient memory utilization.

Q3) Explain why sharing a reentrant module is easier when segmentation is used than when pure paging is used, with an example.

Sharing a reentrant module is generally easier when segmentation is used compared to pure paging. This is because segmentation allows for finer-grained control over memory access & sharing, while pure paging tends to be less flexible in this regard.

Scenario → Consider a multi-user operating system where multiple processes need to execute a common system library or shared code module that is reentrant. This shared code module is used by various processes simultaneously.

Segmentation approach:-

Segmentation divides memory logically - In Segmentation, different parts of the address space are divided into segments, each with

it's own base address & size. you have a dedicated segment for the Shared Code module.

### Shared Code Segment

Create a Shared Code Segment specifically for the reentrant module. All processes that need to access this module reference the same Shared Code Segment in their Segment tables.

### Memory Protection

Can set appropriate access permissions for the Shared Code Segment to ensure that multiple processes can read & execute the code simultaneously, but they can not modify it. This is essential for security in reentrant modules.

**Efficient Sharing -** Since all processes refer to the same Shared Code Segment, there is no duplication of the code in memory. This makes efficient use of memory & any updates to the module are immediately visible to all processes that use it.

### Pure Paging approach:

Paging divides memory into fixed-size pages. Memory is divided into fixed size pages, & each page may contain different parts of code & data.

**Shared Code in page -** To share the reentrant module in a paging system, might need to divide the module into pages & allocate these pages to each process. However this can be inefficient.

Complex, especially if the module spans multiple pages.

**Duplication:** Since each process gets its own copy of the pages containing the reentrant module, this can lead to memory duplication. Multiple instances of the same code module occupy separate memory pages, leading to increased memory usage.

**Synchronization:** Ensuring safe concurrent execution of the shared module might require additional synchronization mechanisms to prevent race conditions & data corruption, as processes may execute their own copies of the module.

Sharing a reentrant module is easier & more efficient with segmentation, because it allows for the creation of a dedicated code segment of efficient memory utilization & fine-grained control over access permissions.

4) Explain in detail about paging in 32-bit & 64-bit architecture.

#### • Address Space

- \* The CPU generates 32-bit virtual addresses.
- \* There are  $2^{32}$  (4GB) possible virtual addresses.

- Page Size

- \* Typical page size in 32 bit systems are 4KB (4096 bytes)
- \* The virtual address space is divided into pages & the physical memory is divided into pages frames of the same size.

- Page Table

- \* To map Virtual addresses to physical addresses a page table is used.
- \* Page table typically consists of multiple entries one for each page in the virtual address space.
- \* Each entry contains the mapping information for a virtual page to a physical page frame.
- \* The page table itself is stored in physical memory.

- Translation Process

To translate virtual address to physical address, the CPU splits the virtual address into a page number & an offset within the page.

It looks up the page number in the page table to find the corresponding page frame number.

The offset within the page is added to the base address of the page frame to form the physical address.

- Memory Protection

\* Paging allows for memory protection by setting access permission on a per-page basis in the page table of entries.

\* Unauthorized accesses result in page faults & are handled by the operating system.

~~05~~ 04

04

## Paging in 64-bit architecture

- Address Space

- In 64-bit architecture, the CPU can generate 64-bit virtual address.

- Page Size.

- Page size in 64-bit systems are typically still 4KB but can be larger, such as 2MB or 1GB.

- Larger page sizes can reduce the overhead of managing a large number of small pages.

- Page Table

- With the vastly expanded address space, the page table becomes more complex.

- A straight forward flat page table with a 1-to-1 mapping of virtual pages to physical frames would be impractical.

- To manage this, hierarchical or multi-level page tables are often used.

- In these systems, the page table is divided into multiple levels, & each level maps a portion of the virtual address space to physical frames.

- Translation process

The translation process in 64 bit system is similar to 32 bit system, but involves traversing multiple levels of page tables to find the physical frame number. This hierarchical approach helps manage the vast address space more efficiently.

- Memory Protection.

remain a key feature in 64-bit systems, allowing for fine-grained control of permissions on a per-page basis, just like 32-bit systems.

- Large page

64 bit architecture often supports large page sizes for improved performance in certain scenarios. Large pages can reduce the size of the page table & improve efficiency.

Q5)

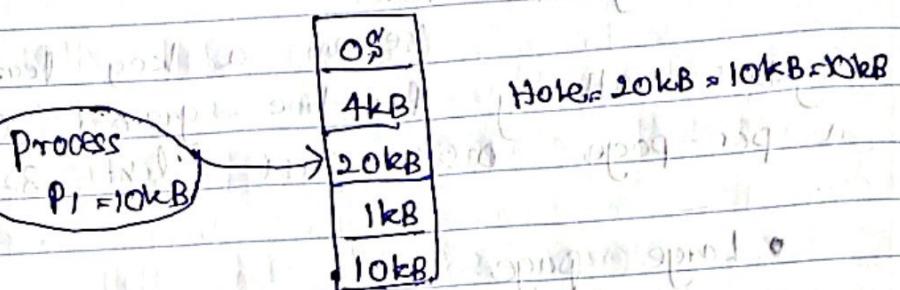
Discuss the given memory management techniques with diagrams.

a) Partition Allocation Methods.

Here, a partition needs to be chosen when there are multiple partitions that are readily available to satisfy a process need. An allocation mechanism for partitions is required in order to select a specific partition. If a partition allocation technique prevents internal fragmentation, it's regarded as superior. Various placement algorithms exist.

First fit:

The partition which is the first free block from the top of main memory is allocated. It starts scanning memory at the beginning & modifies the first sufficiently big block that becomes accessible. As a result, it allocates the first sufficiently huge hole.

2. Best fit:

Assign the process to the partition that among the free available partitions is the smallest sufficient partition in order of precedence. It looks through the complete list of holes to locate the smallest one whose size is either larger than or equal to the processes size.

Process  $P_1 = 10\text{KB}$

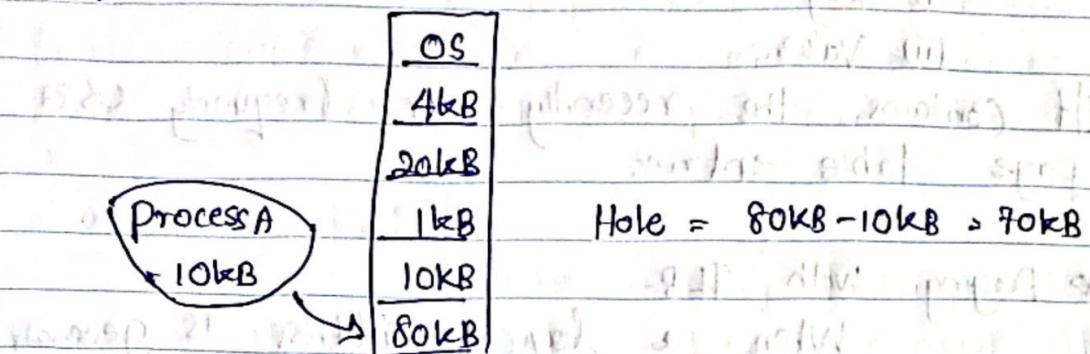
0KB
4KB
20KB
1KB
10KB

Hole:  $10\text{KB} - 10\text{KB} = 0$

3. Worst fit:

Assign the process to the primary memory partition that has biggest suitable size among all the free

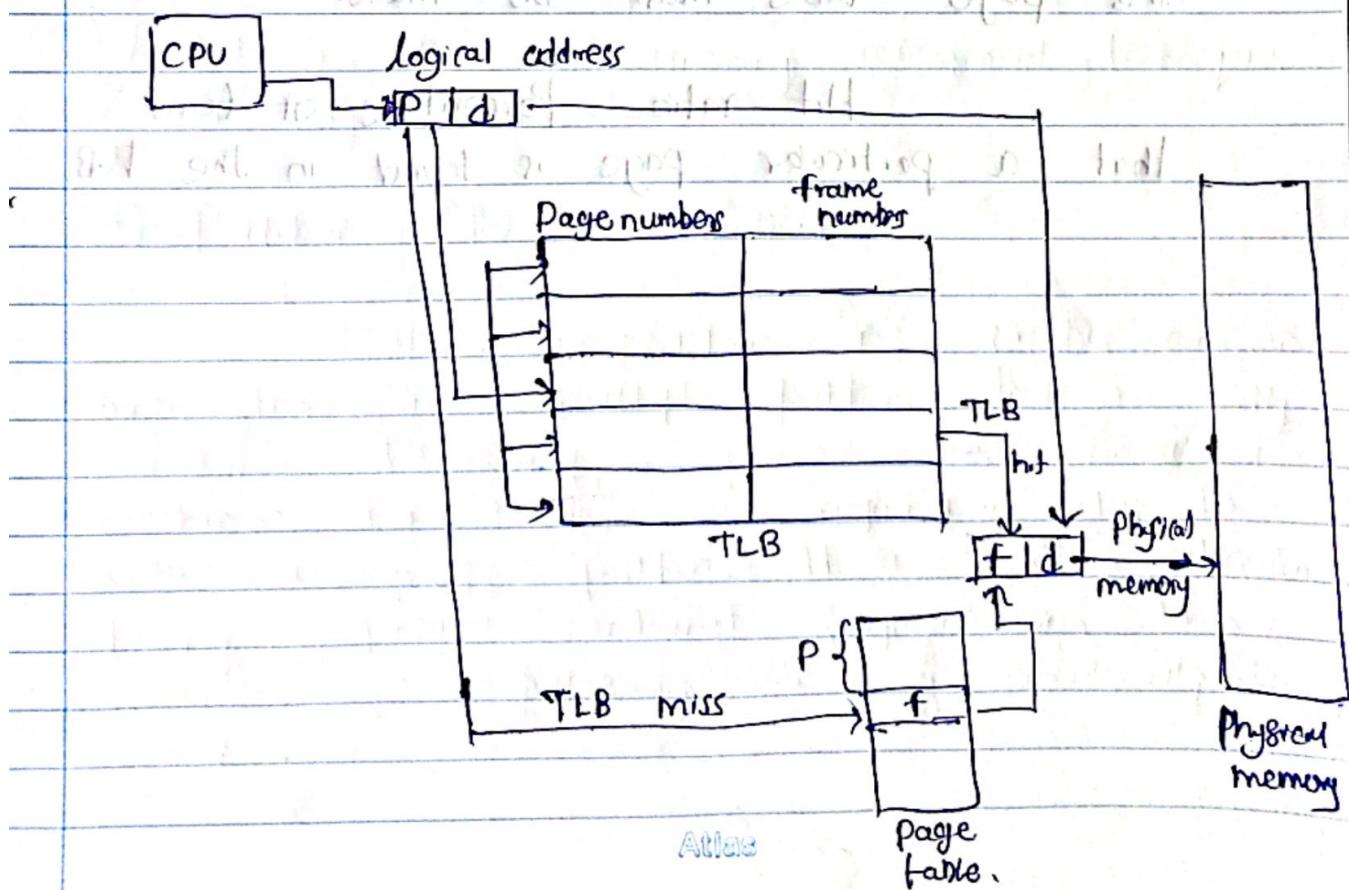
Ponditron. The best fit algorithm is in opposition to it. To identify the biggest hole & assign it to be processed , it Scans the Complete list of holes.



#### 4. Next fit :-

This is similar to the first-fit , but it will search for the first position that is good enough from the last allocation point.

#### b) Paging & Translation Look-aside Buffer.



It is fast lookup hardware cache.

More expensive than main memory.

It has 2 parts:

(i) key

(ii) Value

If contains the recently or frequently used page table entries.

$\rightarrow$  Paging with TLB - When a logical address is generated by CPU, its page number is presented to TLB.

Did it get a TLB hit? If the number is found, its frame number is immediately available & is used to access memory.

TLB miss if the page number is not in the TLB, a reference to the page table must be made.

Hit ratio - Percentage of time that a particular page is found in the TLB.